SNIA CSTI | CLOUD STORAGE TECHNOLOGIES

Ceph Storage in a World of AI/ML Workloads

Live Webinar

January 30, 2025

10:00 am PT / 1:00 pm ET

Today's Presenters

- Michael Hoard
- SNIA
  - CST Chair

- Kyle Bader
- IBM Storage
  - Principal Portfolio Architect

- Phil Williams
- Canonical
  - Product Manager

The SNIA Community

200
Corporations, universities, startups, and individuals

2,000
Active contributing members

50,000
Worldwide IT end users and professionals

Cloud Storage Technologies (CST) Community

Committed to the adoption, growth and standardization of **intelligent data storage usage** in cloud infrastructures.

NEW SNIA Cloud Object Storage Plugfest

This encompasses data services, orchestration and management, as well as the promotion of portability of data in multi-cloud and hybrid cloud environments.

NEW SNIA Cloud Object Storage Test Tools

Learn more at snia.org/cloud

SNIA Legal Notice

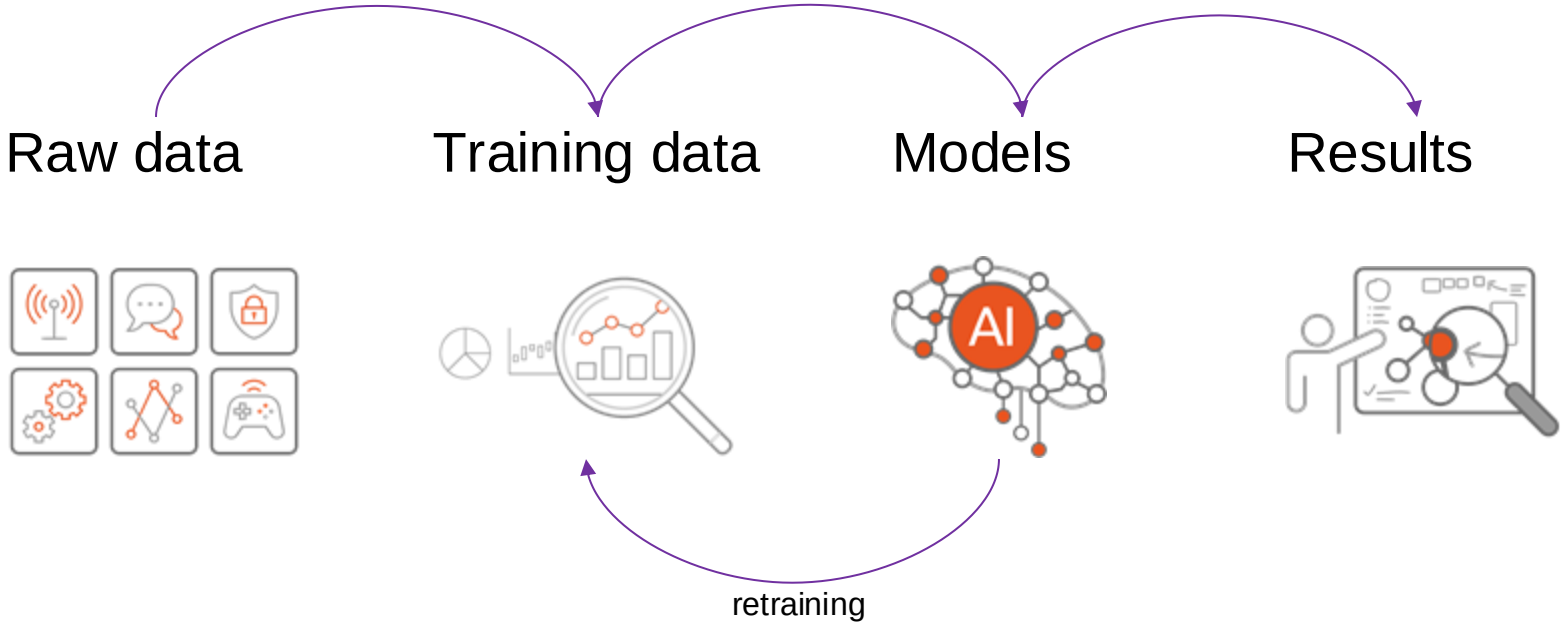
- The material contained in this presentation is copyrighted by SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced in their entirety without modification
  - SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of SNIA.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

  NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.

Agenda

- AI workloads and lifecycle
- Performance needs of Training, Checkpointing and Inference
- Importance of storage in AI infrastructure
- Why Ceph?
- Increasing storage efficiency
- Use cases
- Find out more

AI Workloads / Lifecycle

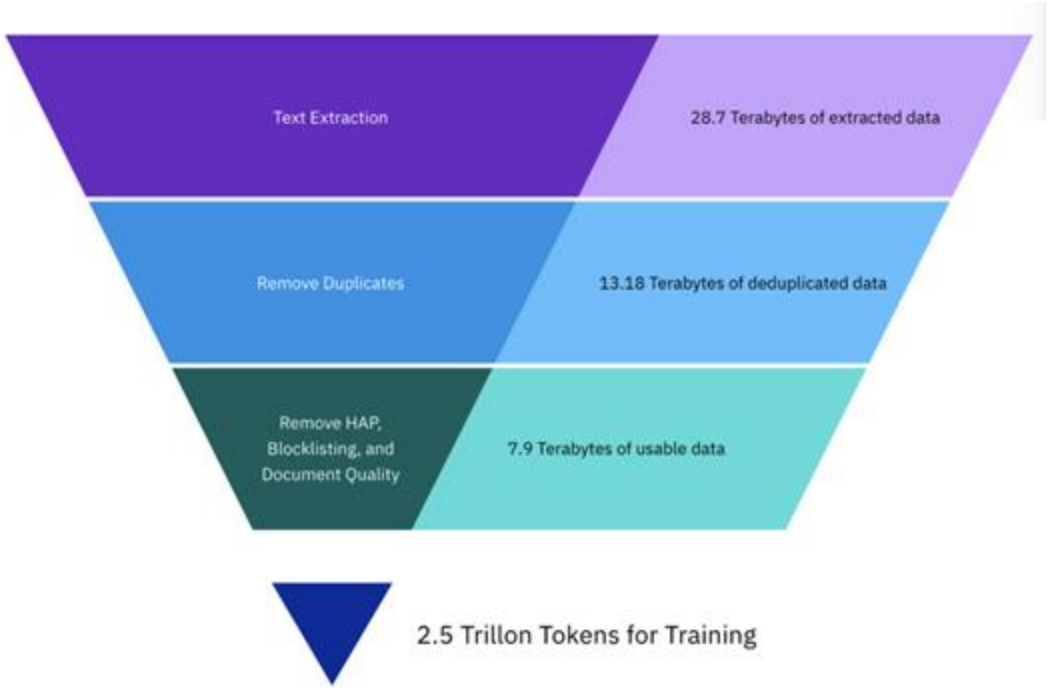


Training

- Usually limited by
  - Network Bandwidth
  - Pre-processing
  - Model architecture

- Typical GPU
  - Upto 4 petaFLOPs (FP8)
  - 5 GBps storage throughput recommended
  - 20 GB/s recommended per reference system

LLMs: Granite 13b Data Pile



GPT3
45TB

570GB

300 billion tokens

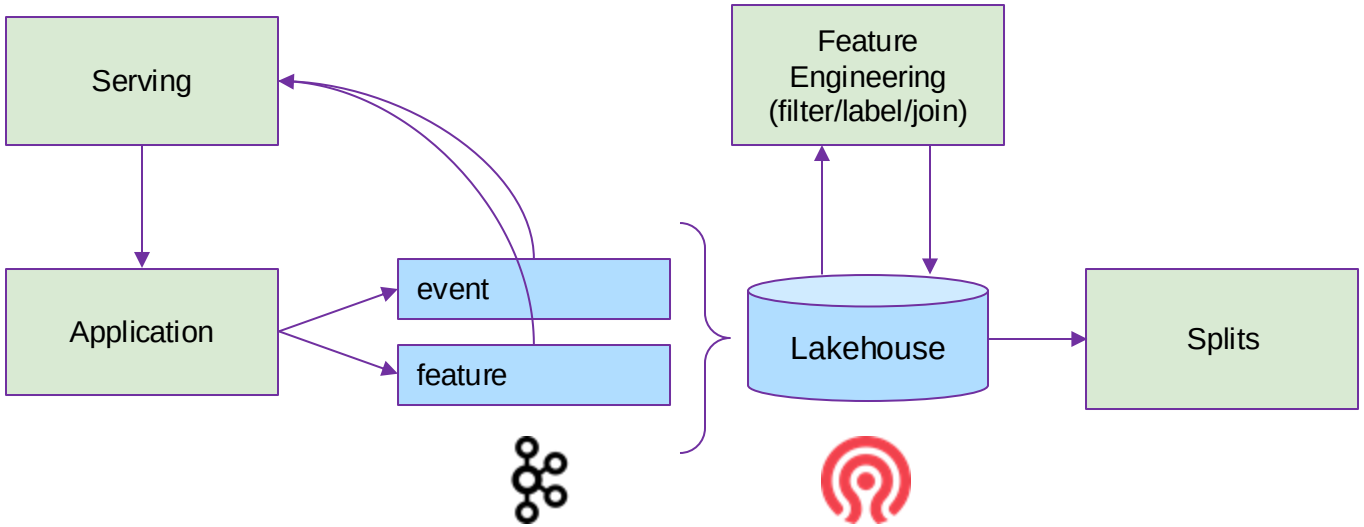
[Granite Foundation Models]
[The Pile: An 800GB Dataset of Diverse Text for Language Modeling]

Checkpointing

Model	Size	Estimated Checkpoint (GiB)	@ 35GB/s
Granite	13b	170	5s
Llama3	70b	913	28s
GPT3	175b	2282	70s
Llama3	405b	5281	162s
DLRM-2021	1t	13039	400s

Checkpoint size estimates based on use of ADAM optimizer
Reducing checkpointing times
Check-N-Run: a Checkpointing System for Training Deep Learning Recommendation Models

Recommendation Systems

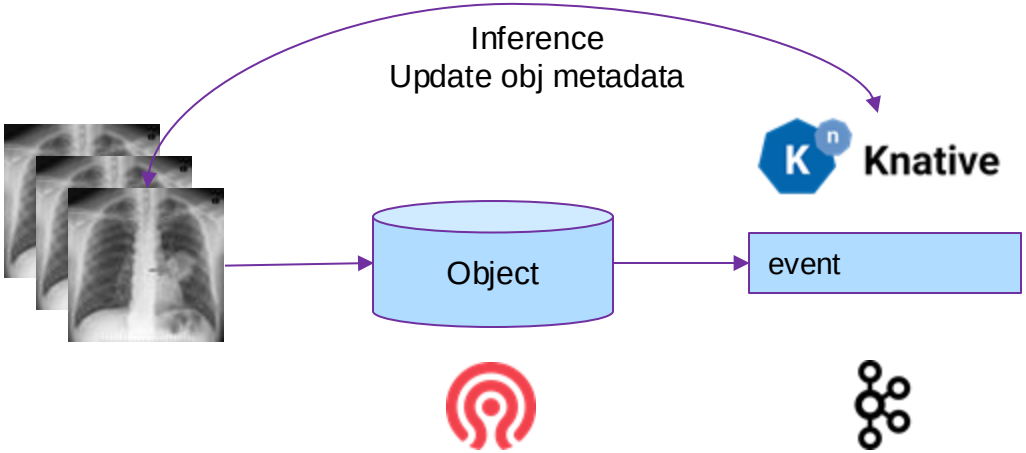


[1] Understanding Data Storage and Ingestion for Large-Scale Deep Recommendation Model Training

[2] Check-N-Run: a Checkpointing System for Training Deep Learning Recommendation Models

"At Facebook's datacenter fleet, for example, deep recommendation models consume more than 80% of the machine learning inference cycles and more than 50% of the training cycles." [2]

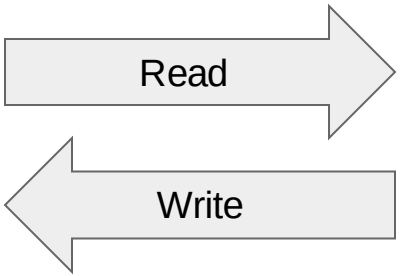
Event Driven Inference



X-Ray Analysis Automated Pipeline

Why is Storage Important?

Dataset



Data processing

Why is Storage Important?
Storage economics

Performance

- App expectations
- Business value

Capacity

- Storage needs only increase over time

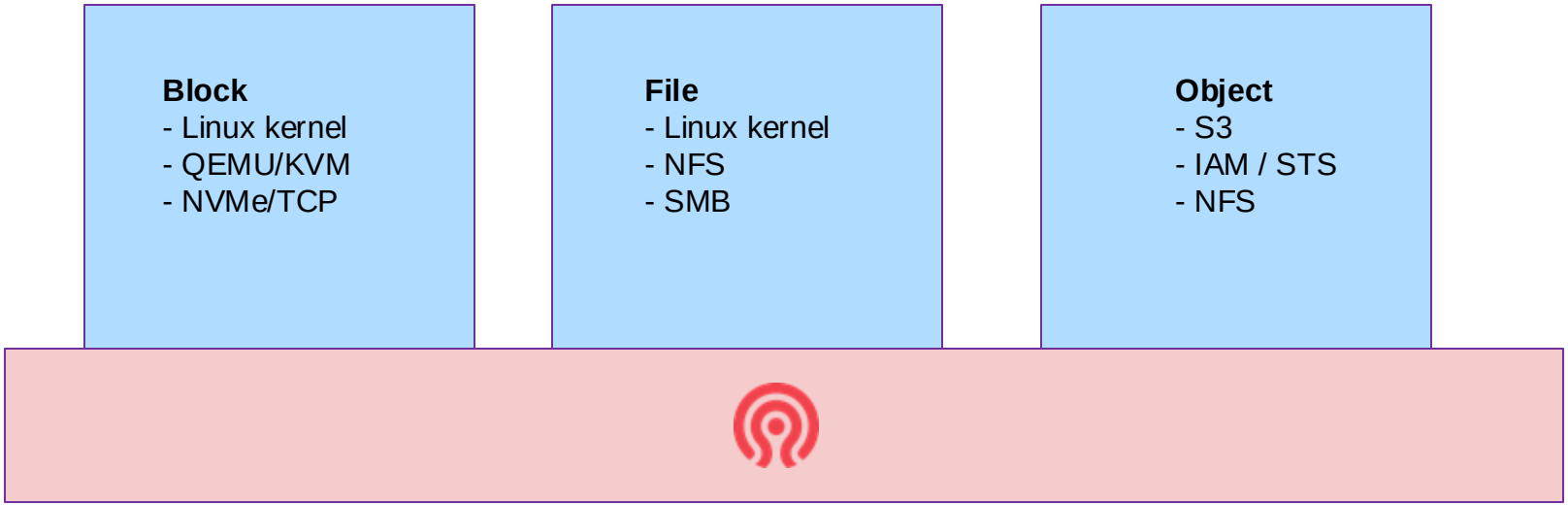
Reliability

- Data cannot be lost
- Data must be available 24/7/365

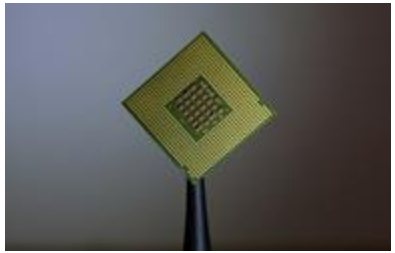
Cost

- Shrinking budgets
- Rising costs

Why Ceph?
Multi-protocol by default



Why Ceph?
Hardware agnostic



CPU and Memory

Higher clock speeds

RAM for cluster ops

Network

High-bandwidth

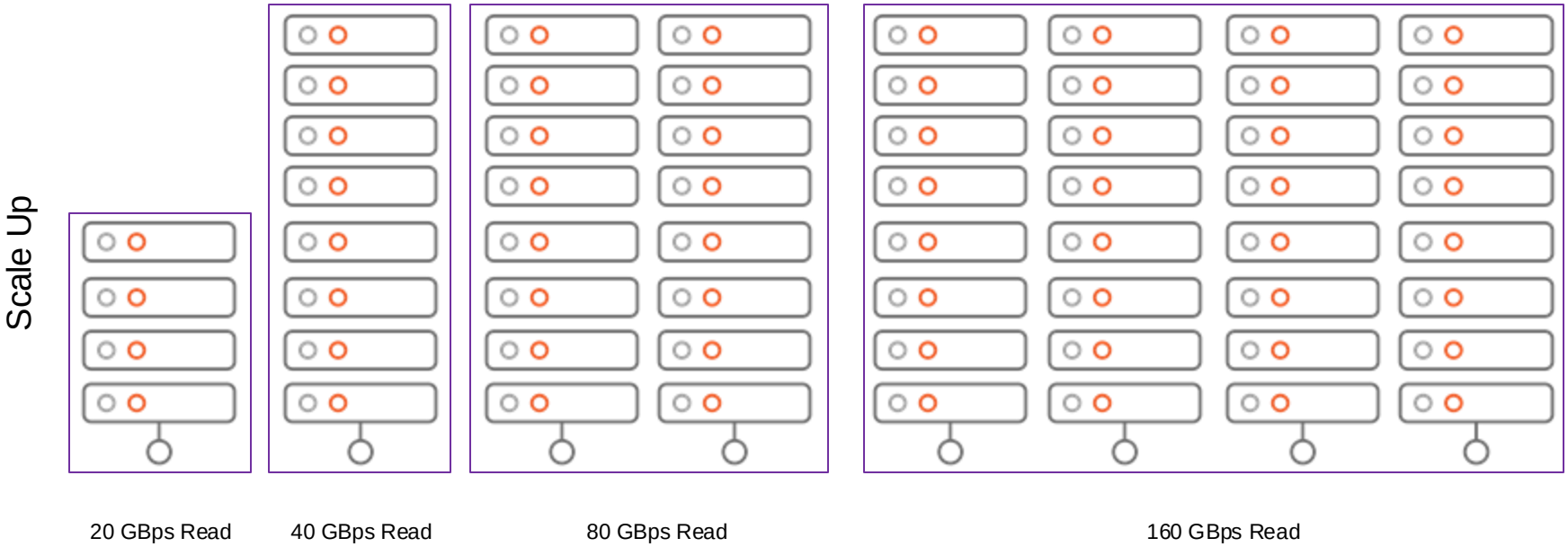
Low-latency

Media

Capacity

Performance

Why Ceph?
Scale from a few nodes to hundreds



Learn more: https://ceph.io/en/news/blog/2024/ceph-a-journey-to-1tibps/

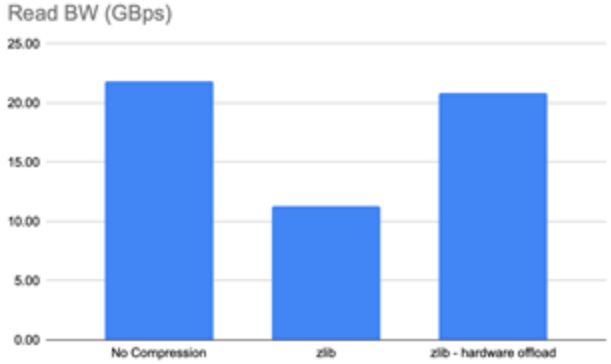
Accessibility

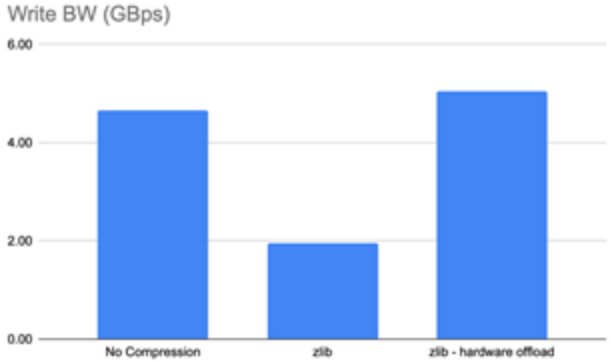
- Availability now
  - Proven in production
- Fully open source
  - Download from https://ceph.io
  - Source at https://github.com/ceph
  - Docs at https://docs.ceph.com/en/squid/
- Support
  - Community mailing list https://ceph.io/en/community/connect/
  - Supported by a wide ecosystem of vendors and practitioners
- No speciality system for AI needed
  - Just the correct planning and design

Increasing Storage Efficiency

- Compression
  - Reduces TCO
  - Applies to all storage media
  - Can lead to CPU overheads
  - Negatively affecting performance

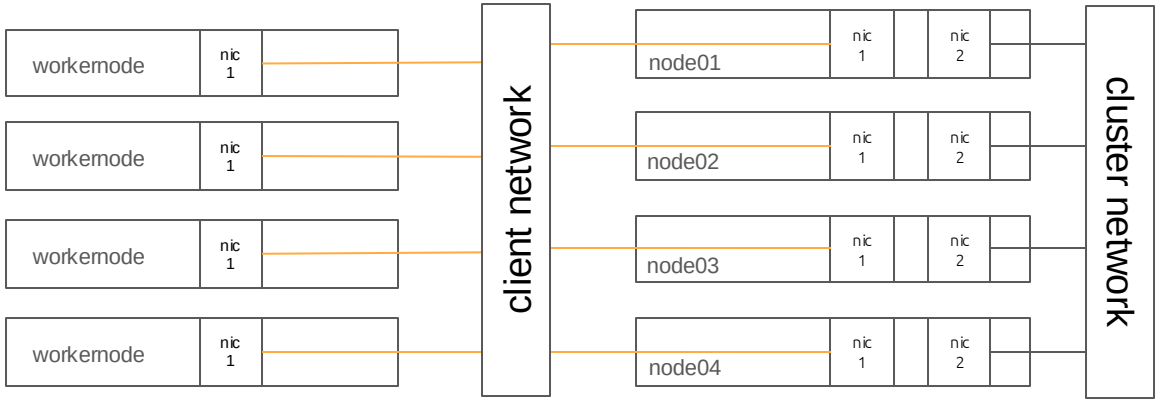
- Hardware Accelerators
  - On-die or PCIe add-in cards:
    - Compression (RGW and Bluestore)
    - SSL
  - Ceph S3 object compression
    - Increase >250% write throughput
    - Increase >180% read throughput
  - Minimal additional hardware cost





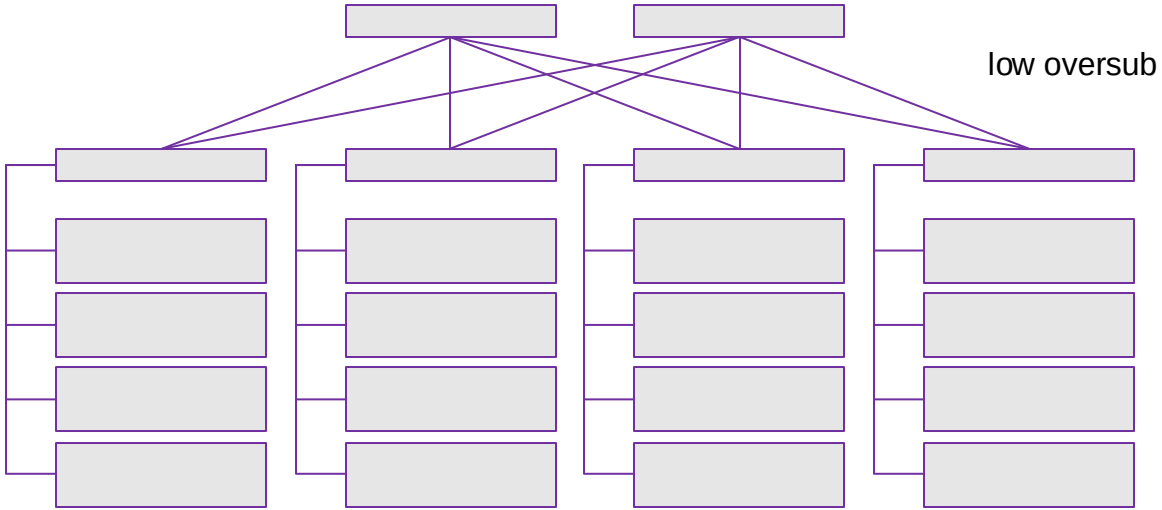
Example Use Cases with Ceph

- Ceph
  - 4 Nodes
    - 2x CPU (32 core ea.)
    - 512GB RAM
    - 2x 100GbE
    - 24x TLC RI NVMe
- Workers
  - 4x GPUs



- Ave Read 30 GB/s
- Ave Write 4.66 GBps

Network



Where to Learn More About Ceph?

- Ceph.io
- Ceph Days
  - Bengaluru, India - 23rd Jan 2025
  - San Jose, USA - 25th March 2025
  - London, UK - 4th June 2025
- Cephalocon
  - 2024 hosted by CERN in Geneva, Switzerland - slides and recordings
  - 2025 TBA
- SNIA Educational Library
  - Ceph: The Linux of Storage Today

Takeaways

- GPUs are expensive, high utilisation is paramount for reducing TCO
- Ceph's approach to scaling helps meet growth demands
- Network planning is key to scaling out
- Hardware agnostic Ceph provides flexibility
- Pluggable architecture allows for integration with hardware offload(s)

Q&A

Thanks for Viewing this Webinar

- Please rate this presentation and provide us with feedback
- This webinar and a copy of the slides are available at the SNIA Educational Library https://www.snia.org/educational-library
- A Q&A from this webinar will be posted to the SNIA Cloud blog: www.sniacloud.com/
- Follow us @SNIACloud

SNIA CSTI | CLOUD STORAGE TECHNOLOGIES

Thank you